Exhibit D

VOICE CONTROL OF MOBILE TELEPHONES

DR. R. EUGENE HELMS

DIRECTOR - PRODUCT DEVELOPMENT
VOICE CONTROL SYSTEMS
16610 DALLAS PARKWAY
DALLAS, TEXAS 75248

SUMMARY

The recent advent of cellular mobile telephone systems has marked the beginning of a new era in telecommunications. Accompanying this significant technological advance has been a proliferation of mobile telephone users, each of whom faces the challenge of manipulating their phones in a mobile environment. Voice control offers the best solution for safe and convenient use of mobile telephones. The premier voice-controlled cellular telephone product, offered by FCA Telecom Ltd., is described in this paper. This product is based on Voice Control Systems' unique speaker-independent voice recognition technology. Also presented is a peek at the future of voice control in mobile telecommunications.

INTRODUCTION

Voice control systems in mobile telecommunications are examined in this paper from the perspective of their technological and functional capabilities. An overview of automatic voice recognition technologies is first presented to highlight technology characteristics which are pertinent to mobile telecommunications applications. Then the focus is turned to voice control of mobile telephones and the premier voice-controlled cellular telephone product offered by FCA Telecom Ltd. This product is described in terms of its functionality and market impact. Finally, future directions for voice control in mobile telecommunications are previewed with an eye toward new products and services.

Today's voice control systems are a product of major advances over the past few years in voice analysis techniques and in microprocessor technology. Figure 1 presents a block diagram of a mobile communications voice control system. Highly sophisticated analyses of digitized speech signals can now be performed using digital circuitry comprised of only a handful of integrated circuits. Semiconductor suppliers continue to make major strides in producing microprocessors and microcomputers of ever-increasing capability and speed. Intel's 80186/286/386 family of microprocessors and Texas Instruments' TMS320 family of digital signal processors are good examples of advanced general-purpose and special-purpose microprocessors, respectively. While today's voice control systems have depended on VLSI advances for their practical implementation, it is the voice processing techniques employed by them that have made systems with such capability possible. These processing techniques stem from recent significant research and development advances in voice control technologies.

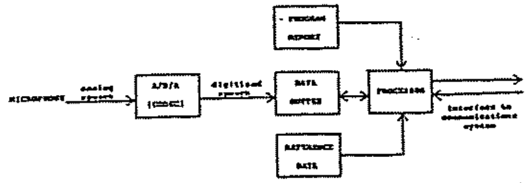

FIGURE 1. BLOCK DIAGRAM OF A MOBILE COMMUNICATIONS VOICE CONTROL SYSTEM

VOICE CONTROL TECHNOLOGIES

Automatic voice recognition systems can be characterized by several dimensions. Among these are system functionality, technology characteristics and system performance. In this section, aspects of these characteristics that have a bearing on the role of voice control in mobile telecommunications are explored. The term "voice control" refers to interactive control through voice commands. As such, a distinction is made here between data entry applications of automatic voice recognition and those that relate to system control.

There are several principal components of voice control system functionality, as depicted in Figure 2, that are pertinent to examine. First, there are vocabulary issues. Automatic voice recognition systems accommodate a finite number of words or phrases at one time. Hence, the user of the system is restricted to those words or phrases that are resident in the system's vocabulary. While vocabulary size requirements vary from one application to another, voice control applications typically require a rather limited vocabulary set. Mechanisms that require interactive control tend to have a somewhat limited scope of control options. This is true, for example, for a driver of a motor vehicle who is controlling a cellular telephone through voice commands rather than tactile interface. By using a voice command syntax structure, similar to a command menu structure common in man-computer interfaces, large command vocabularies can be accommodated in voice control applications by partitioning the total vocabulary into a set of smaller vocabularies. Each such subvocabulary would then correspond to nodes in the syntax tree. This approach is appealing because it retains simplicity for the user through a

SS 007807

DEPOSITION EXHIBIT 6 FOSTER 1/20/05

logical structuring of the commands while effectively making the voice recognition task one of recognizing a number of smaller vocabularies rather than one large vocabulary. The latter point has strong implications for the accuracy of the voice recognizer, as discussed later when system performance issues are addressed.

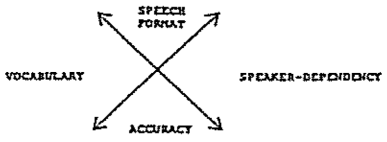

FIGURE 1. FUNCTIONALITY COMPONENTS OF VOICE CONTROL SYSTEMS

Another system functionality component deals with the temporal format of the vocabulary elements that can be accommodated by the voice recognizer. It is very common today for voice recognizers to operate only on discrete speech utterances. The user of such a system is required to pace his input so that the voice recognizer can effectively recognize each vocabulary element prior to receiving a subsequent input utterance. In most cases, this means that the user will pause briefly between inputs to await system response. Such discrete utterance recognition is in contrast with a continuous utterance recognition capability. A continuous speech recognizer does not restrict the user to a single vocabulary element to be input at a time. Rather, the user may string vocabulary words together in a connected or continuous fashion. In so doing, the user is not paced by the recognition speed of the recognizer. The processing power required to implement continuous speech recognizers typically is greater than that required for discrete utterance recognizers. This is due in part to the fact that continuous speech recognizers must make their recognition decisions concurrent with the processing of subsequent inputs. Also, a further level of processing is often required by continuous speech recognizers to handle coarticulation effects. The acoustic variation of utterances spoken in a continuous fashion is greater than when they are spoken discretely. This is attributable to the coarticulation of neighboring sounds wherein the acoustic characteristics of a sound are dependent on those of the previous and following sounds, due to inertia of the physical articulators comprised of the tongue, jaw, teeth and lips. The phenomenon of coarticulation is well-known to speech researchers and presents one of the principal challenges in continuous speech recognizer design.

A third functionality characteristic of automatic voice recognition systems relates to the speaker population accommodated by the voice recognizer. Voice recognizers can be classified as either speaker-dependent or speaker-independent. The vast majority of today's voice recognition systems are of the speaker-dependent variety. These systems are tailored for the voice characteristics of individual users through a training process during which the speaker effectively enrolls his voice in the voice recognition system. This involves uttering each element in the vocabulary several times in order to form reference patterns for that particular speaker. After the user completes the system training procedure, the recognizer will then accommodate that user for that vocabulary. It is necessary for any user of a speaker-dependent system to train the system for their voice before using it in order for the system to achieve acceptable accuracy levels. In contrast, speaker-independent recognition systems do not require system training for individual users. The voice recognition technology embodied in such systems is designed to recognize specific vocabularies for a broad population of users. This is accomplished through a global training procedure for the vocabularies, which is typically performed by the voice recognition system vendor at his facility. The speaker-dependency aspect of system functionality is highly pertinent to mobile telecommunications systems and warrants a close look.

As mentioned previously, voice recognition systems based on speaker-dependent recognition technology are far more prevalent today than are speaker-independent systems. This is due in large measure to the fact that an accurate speaker-independent capability is very difficult to achieve. Clearly, by tailoring the voice recognizer for an individual's voice characteristics, as in the case of speaker-dependent technology, the task of recognizing speech is greatly simplified relative to accommodating everyone's voice characteristics at once. Hence, most recognizers today are speaker-dependent in nature, employing some form of pattern matching wherein the reference patterns gathered by the recognizer during the enrollment process are matched against input speech to make a recognition decision. The inherent advantage in speaker-dependent systems relates to the flexibility that the user has with respect to the vocabulary elements. It is up to the users of such systems to define the vocabulary elements and then enroll their patterns. In environments where vocabulary elements are highly dynamic, speaker-dependent systems offer flexibility in this area. To change vocabulary words, a user merely retrains the system for the new vocabulary set.

This contrasts with speaker-independent systems where the vocabulary typically is fixed by the system vendor. This inflexibility of vocabulary for speaker-independent systems poses little limitation to voice control systems where the interaction structure with the specific mechanism to be controlled is nominally static. The attribute of being independent of speaker proves to be an overwhelming advantage for speaker-independent systems in the voice control arena. The simplicity of use inherent in such systems is significant. The system's ability to be used immediately

SS 007809

by anyone, including those with no previous exposure to the voice recognizer, is a particular advantage in telecommunications applications.

The training procedure necessary for speaker-independent technology, not surprisingly, is more complex than that for speaker-dependent technologies. In order to accommodate the high degree of variation in speech patterns for a large population, highly sophisticated training procedures are required typically. Such speech variations are due to differences in vocal tract characteristics, dialect, accent and personal speaking habits. Also, intra-speaker variations due to individuals' changing speech patterns must be accommodated. The speaker-independent training procedure is often highly data and processing intensive and hence is limited to being performed in the vendor's laboratory for specific vocabulary sets. Training the system for a given vocabulary entails analyzing a large amount of speech data usually in digital form on computer systems. This speech database is what the speaker-independent training process uses to formulate global reference data so that accurate recognition of the vocabulary spoken by any individual may be achieved. The speech database is often comprised of hundreds of speakers so that a good sampling of the user population is present.

When speaker-independent systems are developed for specific applications, such as mobile telephone systems, not only is it important that a good sampling of speaker variations be represented in the database, but it is also important that a comprehensive representation of ambient noise characteristics be present as well. By using application-specific data for the vocabulary training, speaker-independent systems can achieve a high degree of immunity to noise, which is of paramount importance in the mobile vehicle environment. It is through development of such specialized voice control systems that speaker-independent voice recognition can be effectively applied in mobile telecommunications systems by affording a potent combination of high recognition accuracy with noise immunity.

U.S. Gov't

Accuracy of any automatic voice recognizer can be characterized predominately by two types of error rate. The first type of error, and typically the more severe, is termed a "substitution error." A substitution occurs when the recognizer misclassifies one of the vocabulary words. In this case, the recognizer does not know that an error was made, and hence the user must take steps to correct the error. The other major error type is termed a "rejection error." A rejection occurs if the recognizer fails to classify a vocabulary word. When a rejection occurs, the voice recognizer is aware that a problem exists and simply instructs the user to repeat the utterance since it could not make a clear classification decision. This type of error usually is not as critical as a substitution, and the rejection error rate is somewhat greater than the substitution error rate in most voice recognizers. Both of these error rates can be expected to exhibit a decrease as a user becomes more accustomed to using the system.

Technically, a third kind of error can occur, though rates for it are difficult to quantify. This error is termed a "spurious response error." A spurious response is said to have occurred when the recognizer mistakes a sound that is not in the vocabulary for one of the vocabulary words. Such sources for spurious response may be words not in the vocabulary, background speech or even non-speech acoustic events, such as doors slamming, horns sounding or any other spurious background noise. Ideally, an automatic voice recognizer would reject all such spurious sounds. However, automatic voice recognition is not immune to such input. Steps can be taken in the system design that relate to microphone activation and placement to minimize spurious responses. Through appropriate system design, high recognition accuracy along with insensitivity to noise has been achieved for speaker-independent voice control systems for cellular mobile telephones, as described in the next section.

VOICE-CONTROLLED CELLULAR TELEPHONES

What is it in general about voice control technology that makes it appealing for application to mobile telephones? Does voice control offer special advantages to vehicle drivers in using their mobile telephones? Indeed it does! — because voice control is unique. It is unique in that it is based on humans' most natural communication modality. It is the most familiar and convenient mechanism for humans to communicate and hence offers special advantages that relate to interacting with machines. First of all, it requires no inherent training of the user. The user need only be instructed to confine his utterances to those in the voice recognizer's vocabulary. Secondly, voice input allows multimodal communication. It may be combined with other input modalities so that, for example, a tactile interaction might take place simultaneously with the voice commands, providing more efficient control with potential redundancy.

For application to mobile telephones, additional features of voice communication are pertinent. Among these are its ability to be efficiently employed in darkness and at a distance from the mechanism being controlled. Further, it requires no substantial panel space or display area and it permits simultaneous use of hands and eyes for different tasks. These aspects of voice control have a particular significance in the mobile telecommunications environment because drivers are often faced with the challenge of manipulating their telephone equipment while their hands and eyes are busy maneuvering the vehicle. Hence, in addition to increased functionality and ease of operation there are strong advantages

that relate
the use
telephones.

While
been around
advent of c
has marked
telecommunic
significant
proliferatio
each of
manipulating
environment.
offers the
convenient
The complex
phones, par
many specia
demands cor
vehicle dri
control for
highlighted
driver to
road. Cell
control need
control all
Voice contro
ions for the
cellular ph
voice contro
are apparent

In ear
announcement
controlled
made. Brit
distributor
FCA Telecom
voice recog
Voice Contro
produce and
Telephone pi
products wil
cellular m
mounted in
enables upgr
phones to
controlled.
recognition
individual's
modifying th
compartment.
telephone is
that now th
phone by voi

There
characteristi
products. T
is totally
immediate a
voice contr
eliminating
system trai
specialized
specifically
of the mobi
It is a sof
a single
(Intel 80186
circuit inte
through the
control unit
compartment)

SS 007810

error rate substitution izers. Both expected to ecomes more

f error can lifficult to a "spurious response is recognizer ot in the lary words. nse may be background tic events, sounding or nd noise. nizer would . However, t immune to en in the microphone minimize appropriate n accuracy e has been ent voice ar mobile xt section.

PHONES

bout voice : appealing oned? Does antages to ir mobile -- because unique in st natural the most for humans s special cting with equires no The user onfine his ice recog- oice input It may be lities so interaction with the efficient

telephones, munication ts ability rkness and ism being uires no y area and hands and aspects of gnificance nvironment with the telephone eyes are Hence, in y and ease advantages

that relate to safety that are afforded by the use of voice to control mobile telephones.

While mobile telephone systems have been around for many years, the recent advent of cellular mobile telephone systems has marked the beginning of a new era in telecommunications. Accompanying this significant technological advance has been a proliferation of mobile telephone users, each of whom faces the challenge of manipulating their phones in a mobile environment. Voice control technology offers the best solution for safe and convenient use of these mobile telephones. The complexity of use of cellular telephones, particularly with respect to the many special features they offer today, demands considerable attention from the vehicle driver. The utility of voice control for mobile cellular telephones is highlighted by its capability to allow the driver to maintain eye contact with the road. Cellular phones equipped with voice control need only be spoken to in order to control all features of the cellular phone. Voice control can eliminate visual distractions for the driver with regard to using his cellular phone. The safety advantages of voice control of mobile cellular telephones are apparent.

In early January 1986, the premier announcement of the world's first voice-controlled cellular telephone product was made. British Telecom and its largest distributor of cellular telephone products, FCA Telecom Ltd., using speaker-independent voice recognition technology licensed from Voice Control Systems of Dallas, Texas, will produce and market voice-controlled cellular telephone products in early 1986. These products will encompass add-on circuits to cellular mobile telephone transceivers mounted in the trunks of vehicles that enables upgrades of new or existing cellular phones to allow them to be voice-controlled. In this manner the voice recognition feature is added to an individual's phone without physically modifying the telephone set in the passenger compartment. With this enhancement, the telephone is used just as before, except that now the driver merely commands the phone by voice to place the calls.

There are a number of noteworthy characteristics of the British Telecom/FCA products. The voice recognition technology is totally speaker-independent, affording immediate accessibility to the products' voice control features for anyone by eliminating the need for voice recognition system training by each user. It is a specialized recognition system designed specifically for the noise characteristics of the mobile communications environment. It is a software-based recognizer requiring a single general-purpose microprocessor (Intel 80186) to implement. The recognition circuit interfaces to the mobile telephone through the bus that connects the phone control unit (telephone set in the passenger compartment) to the transceiver unit (mounted in the trunk) as depicted in Figure 3. The voice control unit may be added to installed cellular telephones or may be acquired with the phone at the time of purchase and installed along with the cellular phone. In either case the voice control unit provides for easy interface with the phone.

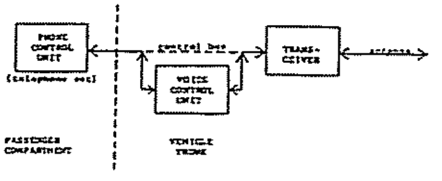

FIGURE 3. BLOCK DIAGRAM OF VOICE-CONTROLLED MOBILE CELLULAR TELEPHONE

The functional operation of the voice control unit centers around syntactically-structured voice commands from the user and voice responses from the voice control unit. The command syntax structure is shown in Figure 4. The output channel of the CODEC in the voice recognizer front-end is used for voice responses to guide the user and provide aural feedback for validation of input. As shown in Figure 3 above, the voice control unit taps into the voice channel and control interface on the control bus of the cellular telephone system. In so doing, the voice control unit recognizes voice commands input to the phone and then issues appropriate commands to control phone operation. Each input command to the phone is acknowledged by the voice control unit through an aural response. If the command is accepted, a short beep-tone signifies to the user that the voice control unit has recognized the command and is ready for the next command to be input. If no clear recognition decision can be made by the voice control unit, the response "repeat" is made to the user requesting him to to re-enter that command.

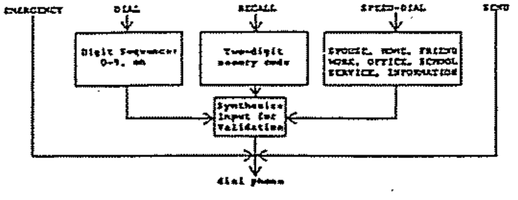

FIGURE 4. VOICE COMMAND STRUCTURE

As indicated in Figure 4 above, five primary commands may be given to the phone. They are DIAL, SPEED-DIAL, RECALL, EMERGENCY, and SEND. The DIAL command is followed by a sequence of discrete digits which is terminated by the command VERIFY, whereupon the digit sequence is repeated to the user for validation. If the user wants the call to then be placed, he says SEND; otherwise he can issue the command CLEAR to effectively restart the command sequence. The command SPEED-DIAL is followed by one of a set of call destination descriptors such as HOME, OFFICE, SCHOOL, etc. Each of these descriptors is associated with a unique

SS 007811

memory location. When the word HOME is recognized, for example, the phone number previously stored in memory location N is accessed, whereupon the voice response unit says "DIALING HOME". Again, if the user desires to place that call he utters the command SEND; otherwise he can say CLEAR to abort the call placement. The command RECALL is functionally the same as SPEED-DIAL except that a two-digit memory location is entered directly following the command. This allows repertory dialing using numeric speed codes. The primary command EMERGENCY will immediately effect call placement to the emergency number in memory without qualifying commands or validation. The remaining primary command is SEND and is used to place a call to the last number dialed, allowing an efficient means for redialing the last called party.

FUTURE DIRECTIONS

There are several exciting directions that voice control in mobile telecommunications will take in the future. First of all, extensions of voice control of cellular phones into the speakerphone realm is imminent. The speakerphone option available with cellular mobile telephones enables the user to carry on phone conversations without having to hold the phone handset. By designing the voice control unit to use the visor-mounted microphone that serves as the input mechanism for the speakerphone option, a totally hands-free capability is provided. The driver merely gives voice commands to U.S. Gov't phone, which are interpreted through the remote microphone on the sun visor, and then proceeds to converse with the called party through the same microphone. Hence, the voice control unit eliminates all visual and tactile distractions for the driver with regard to using his cellular phone.

Another interesting direction for voice control in mobile telecommunications is that of information access. The cellular phone is no less than an information gateway for people in a mobile environment. The capability for a vehicle driver to access databases, voice store-and-forward systems, etc. while he is literally "on the road" is of great value. It does, however, necessitate that the user have easy and efficient access to these information systems. Hence, voice control for information access in the mobile environment is an avenue sure to see substantial product development activity in the very near future.

CONCLUSION

The advancement of automatic voice recognition technology has brought us to the point today that voice control of mobile telecommunications systems is far more than simply an intriguing concept. Practical, cost-effective voice control products for cellular mobile telephones are on the threshold of being shipped to consumers. British Telecom and its affiliate FCA Telecom Ltd. enjoy pioneering positions in this arena and appear to be postured on the forefront of the emerging era of voice-controlled mobile communications equipment.

This paper
equipment fo
developed s
algorithm ha
speech rec
Evaluation t
running car
recognition

1. INTRODU

Recently for
hands-free
(voice acti
enormously.
equipment f
running car
recognit
environment
speaker-de
recognition
attributable
fact that
speech suc
available,
exactly det
the followi

(1) Analys
relatio
analyze
the bas

(2) Word b
statist
and sp
noise,

(3) Matchin
such bo
boundar

The speech
technique
recognitior

2. SPEECH

This speec
parts:
detection
Figure 1.

SS 007812